

Figure 1 (SEQ ID NO 1): Amino acid sequence of the *A. chrysogenum* protein encoded by a nucleic acid molecule of the invention (depicted from the N terminus to the C terminus)

MASPIASAAL KARIRRP SML KKLCKPQDLM HHFPNGSYIG WSGFTGVGY P KKMP TYMADH
VEQNG LQGKL KYS LFVGASS GAETENRWAS LDMIDRRTPH QVGKAISKGI NEGKIHF FDK
HLSMF PVDLV YGYYTKDRPH NKLDVVVVEA TDIKEDGSIV PGASVGATPE LIQMADKIII
EVNTSLPSFE GLHDITMTDL PPLRKPYLVM GVEDRIGRTS IPIDPEKVVG ILES DYQDAT
APNAE ADESA NKIAGHLIEF FEHEVAHGRL PNSLLPLQSG IG NVANAII G GLDNSNFRNL
KVVTEVIQDT FLDLFDSGRL DFATATSIRF SPDGFRRFYD NWEAYYGKLL LRSQQVSN SP
EIIRRLGVIA MNTPVEVDIY AHANSTCVMG SRMLNGLGGS ADFLRSSKYS IMHTPSTRPS
KTDPHGVSCI VPMCTHIDQT EHDLDVIVTE QGLADV RGLS PRERARVIK KCAHPVYQPI
LTHYFEKAES DCLRKGWGHE PHLLFNSFDL HKALVEHGSM QKVGQW

Figure 2 (SEQ ID NO 2): Genomic DNA sequence of the coding region of a genomic clone of the new *A. chrysogenum* gene from the translation start codon (ATG) to the last coding codon (TGG). The introns are depicted underlined. A single strand in 5'-to-3' orientation is shown.

```

ATGGCATCAC CAATAGCCTC TGCCGCCCTC AAGGCGCGCA TTCGCCGCCC CTCGATGCTC
AAGAAGCTGT GCAAGCCCCA GGATTTGATG CATCACTTCC CCAATGGCTC GTACATTGGC
TGGTCCGGCT TCACCGGCGT CGGCTACCCG AAGTGAGTTC CACCGTCATC CCGCTCCACA
GTAGGCGCAG CCGGCCCGCT GACAGTCCCC GACAGGAAAA TGCCGACCTA CATGGCCGAT
CACGTCGAGC AGAACGGCCT TCAGGGCAAG CTGAAGTACT CGCTATTCTG GGGCGCATCG
TCGGGTGCTG AGACAGAGAA TCGCTGGGCG TCGCTCGACA TGATTGATAG GAGGACCCCG
CATCAGGTCG GCAAGGCCAT CTCCAAGGGC ATCAATGAGG GCAAGATCCA CTTCTTCGAC
AAGCATCTCT CCATGTTCCC CGTGGACCTT GTATACGTAC GTCAACGATG ATCCCTTGGA
ATGTGCATGT ACTACGAGTA CCTGGCGCTA ACATCCGGTC AGGGCTACTA CACAAAGGAT
AGACCCACA ACAAGCTGGA CGTGGTGGTG GTGGAGGCCA CCGACATCAA AGAGGACGGA
AGCATTGTAC CCGGAGCTTC AGTCGGCGCG ACCCCCGAGC TCATCCAGAT GGCCGATAAG
GTGAGCAATT TCGATTTCTA GCGGAGGGCG CAGCAGGACC TGACATCTCC CTGTGCAGAT
CATTATCGAG GTCAACACCT CACTGCCTTC ATTTCGAGGGT CTCCACGACA TCACCATGAC
CGACCTGCCC CCGCTACGGA AGCCCTATCT CGTCATGGGT GTCGAGGACC GCATCGGCAG
GACCTCTATC CCTATCGACC CCGAGAAGGT TGTAGGCATC CTCGAATCCG ACTACCAGGA
CGCCACTGCC CCCAACGCCG AGGCCGACGA GAGTGCGAAC AAGATTGCTG GCCACTTGAT
TGAGTTCTTC GAGCACGAGG TCGCCACGG CCGTCTCCCG AACTCCCTCC TTCCCCTCCA
GTCCGGCATC GGCAACGTCG CCAACGCCAT CATCGGTGGC CTCGACAAC CCAACTTCCG
CAACCTCAAG GTCTGGACTG AGGTTATCCA GGACACCTTC CTCGACCTCT TCGACTCGGG
CCGCCTCGAC TTTGCCACGG CCACCTCTAT CCGCTTCTCC CCCGACGGTT TCCGCCGGTT
CTACGACAAC TGGGAGGCCT ACTACGGCAA GCTCCTCCTC CGCAGCCAGC AGGTGTCCAA
CTCGCCCGAG ATCATCCGCC GCCTTGGTGT CATTGCCATG AACACCCCCG TCGAGGTCGA
CATCTACGCC CACGCCAACT CCACCTGCGT CATGGGCTCG CGCATGCTCA ACGGCCTGGG
CGGCTCCGCC GACTTCCTGC GCTCCTCCAA GTACTCTATC ATGCACACCC CGTCCACCCG
CCCCTCCAAG ACCGACCCGC ACGGCGTCTC GTGCATCGTT CCCATGTGCA CCCACATCGA
CCAGACTGAG CACGACCTCG ACGTCATCGT CACCGAGCAG GGCTTGCCG ACGTGCGCGG
CCTGAGCCCC AGGGAGAGGG CCCGCGTCAT CATCAAGAAG TGCGCCACC CGGTCTACCA
GCCCATCCTG ACCCACTACT TTGAGAAGGC CGAGAGCGAC TGCCTACGCA AGGGCTGGGG
CCACGAGCCC CATCTGCTCT TCAACTCGTT TGACCTGCAC AAGGCCCTCG TGGAGCACGG
AAGCATGCAG AAGGTCGGGC AGTGG

```

Figure 3 (SEQ ID NO 3): cDNA sequence of the coding region of the new *A. chrysogenum* gene from the translation start codon (ATG) to the last coding codon (TGG). A single strand in 5'-to-3' orientation is shown.

```
ATGGCATCAC CAATAGCCTC TGCCGCCCTC AAGGCGCGCA TTCGCCGCCC CTCGATGCTC
AAGAAGCTGT GCAAGCCCCA GGATTTGATG CATCACTTCC CCAATGGCTC GTACATTGGC
TGGTCCGGCT TCACCGGCGT CGGCTACCCG AAGAAAATGC CGACCTACAT GGCCGATCAC
GTCGAGCAGA ACGGCCTTCA GGGCAAGCTG AAGTACTCGC TATTCGTGGG CGCATCGTCG
GGTGCTGAGA CAGAGAATCG CTGGGCGTCG CTCGACATGA TTGATAGGAG GACCCCGCAT
CAGGTCGGCA AGGCCATCTC CAAGGGCATC AATGAGGGCA AGATCCACTT CTTGACAAG
CATCTCTCCA TGTTCCTCGT GGACCTTGTA TACGGCTACT ACACAAAGGA TAGACCCAC
AACAAAGCTGG ACGTGGTGGT GGTGGAGGCC ACCGACATCA AAGAGGACGG AAGCATTGTA
CCCGGAGCTT CAGTCGGCGC GACCCCGAG CTCATCCAGA TGGCCGATAA GATCATTATC
GAGGTCAACA CCTCACTGCC TTCATTCGAG GGTCTCCACG ACATCACCAT GACCGACCTG
CCCCCGCTAC GGAAGCCCTA TCTCGTCATG GGTGTCGAGG ACCGCATCGG CAGGACCTCT
ATCCCTATCG ACCCCGAGAA GGTGTAGGC ATCCTCGAAT CCGACTACCA GGACGCCACT
GCCCCAACG CCGAGGCCGA CGAGAGTGCG AACAAGATTG CTGGCCACTT GATTGAGTTC
TTCGAGCACG AGGTCGCCCC CGGCCGTCTC CCGAACTCCC TCCTTCCCCT CCAGTCCGGC
ATCGGCAACG TCGCCAACGC CATCATCGGT GGCCTCGACA ACTCCAACCT CCGCAACCTC
AAGGTCTGGA CTGAGGTTAT CCAGGACACC TTCCTCGACC TCTTCGACTC GGGCCGCCTC
GACTTTGCCA CGGCCACCTC TATCCGCTTC TCCCCGACG GTTTCGGCCG GTTCTACGAC
AACTGGGAGG CCTACTACGG CAAGCTCCTC CTCCGAGCC AGCAGGTGTC CAACTCGCCC
GAGATCATCC GCCGCCTTGG TGTCATTGCC ATGAACACCC CCGTCGAGGT CGACATCTAC
GCCCACGCCA ACTCCACCTG CGTCATGGGC TCGCGCATGC TCAACGGCCT GGGCGGCTCC
GCCGACTTCC TGCCTCCTC CAAGTACTCT ATCATGCACA CCCCCTCCAC CCGCCCCCTC
AAGACCGACC CGCACGGCGT CTCGTGCATC GTTCCCATGT GCACCCACAT CGACCAGACT
GAGCACGACC TCGACGTCAT CGTCACCGAG CAGGGCCTGG CCGACGTGCG CGGCTGAGC
CCCAGGGAGA GGGCCCGCGT CATCATCAAG AAGTGCGCC ACCCGGTCTA CCAGCCCATC
CTGACCCACT ACTTTGAGAA GGCCGAGAGC GACTGCCTAC GCAAGGGCTG GGGCCACGAG
CCCCATCTGC TCTTCAACTC GTTTGACCTG CACAAGGCCC TCGTGGAGCA CGGAAGCATG
CAGAAGGTCG GGCAGTGG
```

Figure 4 (SEQ ID NO 4): Genomic DNA sequence of a BamHI/EcoRI fragment of a genomic clone of the new *A. chrysogenum* gene (a single strand in 5'-to-3' orientation is shown). The translation start codon (ATG) and the translation stop codon (TAA) of the coding region are depicted underlined and in bold type; the introns are depicted underlined.

```
GAAG ATCGCATTTG GCGAGTGGG CTAATAATGC CTGCTGCCTG CCTGTGGACG
GTAAATGAAT TAGGTGGAAT GTGTCGCAA TTAGAGGGGAA TGGCCCCCTT ATCATATAAA
GTGCCAATGC GATACTATGG CGTGGCGTGG GGTCCGCTCG GTGTCCGGCC GGTCGAACGG
AGGTCCCGGC TATCAATAGG CGGTAGGCCG GCATTGAATC GGTTCACCG TATTCCAGAC
ACCCAAGGAA GGCCCGCCAC CCCCAGCTCC GGCCTGGGGA TAGCGCCGAG TGGAGCACTC
ACGGGGGCCG TGTTTGACTC GAAGACGCGT CGTGATTGGC CAGAACTTCA TCCCCCTCTG
CCAAGTATTG GTTCACGGGA TTCGGCGACG TCAACGACCC CACCGGCCCG GATTACATAA
GGTGCACGTC AGCTACTACG TAGTACTCGT ACTTGGAAG GAGGGACCCT TGGGGTCGGA
GGTTTTAAAG GCAATGGCTT CTTGCTGGT CCACCCAACC TGACTCTCAC TCTCCCTTTT
ACCTCGCTCC TCTGATTATT CCCTCGTCTG CGTCTGGATT TCATCTCTTT CCCCTCCCGG
CCCCTTTGGA TCTCTGCTCT CCCCTCCTCT CTCCCCGCA TTGGTGTGTA AAACCACTGT
CCCGCGGCCCT CGCGACGAGT GACGTACTGC AAGCCGAAAC CTCACAATCC CTTCCTCACA
ATGGCATCAC CAATAGCCTC TGCCGCCCTC AAGGCGCGCA TTCGCCGCC CTGATGCTC
AAGAAGCTGT GCAAGCCCCA GGATTTGATG CATCACTTCC CCAATGGCTC GTACATTGGC
TGGTCCGGCT TCACCGGCGT CGGCTACCCG AAGTGAGTTC CACCGTCATC CCGCTCCACA
GTAGGCGCAG CCGGCCCGCT GACAGTCCCC GACAGGAAAA TGCCGACCTA CATGGCCGAT
CACGTCGAGC AGAACGGCCT TCAGGGCAAG CTGAAGTACT CGCTATTCGT GGGCGCATCG
TCGGGTGCTG AGACAGAGAA TCGCTGGGCG TCGCTCGACA TGATTGATAG GAGGACCCCG
CATCAGGTCG GCAAGGCCAT CTCCAAGGGC ATCAATGAGG GCAAGATCCA CTTCTTCGAC
AAGCATCTCT CCATGTTCCC CGTGGACCTT GTATACGTAC GTCAACGATG ATCCCTTGGA
ATGTGCATGT ACTACGAGTA CCTGGCGCTA ACATCCGGTC AGGGCTACTA CACAAAGGAT
AGACCCACA ACAAGCTGGA CGTGGTGGTG GTGGAGGCCA CCGACATCAA AGAGGACGGA
AGCATGTGAC CCGGAGCTTC AGTCGGCGCG ACCCCCGAGC TCATCCAGAT GGCCGATAAG
GTGAGCAATT TCGATTTCTA GCGGAGGGCG CAGCAGGACC TGACATCTCC CTGTGCAGAT
CATTATCGAG GTCAACACCT CACTGCCTTC ATTGAGGGT CTCCACGACA TCACCATGAC
CGACCTGCCC CCGCTACGGA AGCCCTATCT CGTCATGGGT GTCGAGGACC GCATCGGCAG
GACCTCTATC CCTATCGACC CCGAGAAGGT TGTAGGCATC CTCGAATCCG ACTACCAGGA
CGCCACTGCC CCAACGCCG AGGCCGACGA GAGTGCGAAC AAGATTGCTG GCCACTTGAT
TGAGTTCTTC GAGCACGAGG TCGCCACGG CCGTCTCCCG AACTCCCTCC TTCCCCTCCA
GTCCGGCATC GGCAACGTCG CCAACGCCAT CATCGGTGGC CTCGACAAC CCAACTTCCG
CAACCTCAAG GTCTGGA CTG AGGTTATCCA GGACACCTTC CTCGACCTCT TCGACTCGGG
CCGCCTCGAC TTTGCCACGG CCACCTCTAT CCGCTTCTCC CCCGACGGTT TCCGCCGGTT
CTACGACAAC TGGGAGGCCT ACTACGGCAA GCTCCTCCTC CGCAGCCAGC AGGTGTCCAA
CTCGCCCGAG ATCATCCGCC GCCTTGGTGT CATTGCCATG AACACCCCCG TCGAGGTCGA
```

CATCTACGCC CACGCCAACT CCACCTGCGT CATGGGCTCG CGCATGCTCA ACGGCCTGGG
CGGCTCCGCC GACTTCCTGC GCTCCTCCAA GTACTCTATC ATGCACACCC CGTCCACCCG
CCCCTCCAAG ACCGACCCGC ACGGCGTCTC GTGCATCGTT CCCATGTGCA CCCACATCGA
CCAGACTGAG CACGACCTCG ACGTCATCGT CACCGAGCAG GGCCTGGCCG ACGTGCGCGG
CCTGAGCCCC AGGGAGAGGG CCCGCGTCAT CATCAAGAAG TGCGCCACCC CGGTCTACCA
GCCCCATCCTG ACCCACTACT TTGAGAAGGC CGAGAGCGAC TGCCTACGCA AGGGCTGGGG
CCACGAGCCC CATCTGCTCT TCAACTCGTT TGACCTGCAC AAGGCCCTCG TGGAGCACGG
AAGCATGCAG AAGGTCGGGC AGTGGTAAGA TTGGCGAGAC GGGAGAGGCG TTGTTGTAGG
AGTTGGA ACT AGAATCAGAT ATACAGCCTT TCATATATGT AGATAATGGA GCCATT

Figure 5 (SEQ ID NO 5): A. chrysogenum genomic DNA sequence of an approx. 16 kb region marked by SnaBI and BfrI and containing the biosynthesis genes pcbC (position 1366 to position 350, inverse arrangement) and pcbAB (position 2598 to position 13517). A single strand in 5'-to-3' orientation is shown. The particular translation start codons (ATG, GTG) and the particular translation stop codons (TAA, TGA) of the respective coding regions are depicted underlined and in bold type. The said cleavage sites are depicted underlined.

```

1      TACGTACATA CGTCGCGGGG GGTAGACAAT GGTGTGGTGT ACGTGTACAA CTACAGTCAG
61     ACATGGACGC AGGAAACGCA TCATTGATAC ATGCACACGG GGCAGCAAAT TTAGCCTGTT
121    TCACTACATG TACATAGAGG GTACACTCCA GAGCATACTG ATGGGAGAAA AAGGGTTCTGA
181    TTGCTGGTGG TTTAACATAG CCGGCAAGGG GAAAAAAAAA AGGGGGCGGA GAAGGACTGA
241    TTCTTCCTGG CAGACACTCG ACCCTTCCGG CCCCTTGAAC TGCTTTTACT CCCGCATTCC
301    TCCGCACGCC CGCCCACAGC GGCAGATCAG CCGAACCTGA TCGACCGATT TAGGTCTGAC
361    CATTCTTGTT GATCAAGCCC CGCAGTCCCC CCTGCAGATA CTCTCCGTAG GAGATGGCCG
421    GCTTGTCCTT GCGGGCATCC TTGGCCCCAT CCTTGGCGGT CGCGGGGTCC CACGGCTGGA
481    TGGTGTCTTC CCAGCCCAGG TTGACGAAGA AGGGCAGTGA CTGGCGCTCC TCGTTGACCC
541    ATTTGACGCG GTGGATCGGG GCCGGGTAGT AGTCGTCGGT GATATGGGCC ATGTAGCTGC
601    CGCAGTTGAT GAGGAAGCCC GTGTCGTCAG CCTGGATGTC CTGCCAGCCC TCGGGGGTCT
661    TGACCTGCAG ATTCTGCACG TCGGACTGGT ACAACACCGT GATGAGGGAC ACGTCCTCGT
721    GCCACTCGAA GCTGAGCTTG GTGCCGTCTG CGGCCGTCTT GATGGCCGGC TCCGGGTACG
781    GGTGAGGTA CGGGTAACGG ATGAGCACGA CCGACGAGAG CGTCGTGTCA CGGCGGGAGT
841    GCGGGGTGAA GAAGTCCTCG TCGCGACCTA GGGCGAGAGC GTAGCCGCGC AGCACCGCGG
901    AGGAGAGGCC GAAGACGTCC CAGTAGTACT TCTCGGCGAA GGCCCGGAAC CCCGGGTGCT
961    TCGCCTCGTC CGGCCAGACG TTGACCTCGT GCATAGGGGT GGGCTCCTTG ATTCGCGGGT
1021   GGTCTGGGCT GAAGGAGGGG TTCAGGTAGC AGAACGATTC GACCGCCTTC TTGCCCGGGA
1081   TCGGCAGGTA GTAGCCCACC CGGATCTGGG ACTCGTGCTC CTTGTTGTAG GCCCGGATGG
1141   CGAGCTGCCA CTTCTCCTCG TCCGTGATGC TCATGTGGAA TTTGTTTCGT TCGCGCGAGA
1201   GCCACGGCAG GTCGACACCG TGGTTCACCG CGTAAAAGAA GCCTGTGTCTG CGCGATGCGG
1261   CGTCGATGGC GCGAGCTACC TCGAGCTTCT TCTCCTTGTC ATCGCCGAAT AGGGGCGAGA
1321   CATCGATTCT GGGGACGTTG GCCACTGGAA CTGGAACGGA ACCCATGGTG ACGGTTTGTCT
1381   CTGCCTGGTG TAAGATGTGA AAGACGAGAT ATGCGTGAGT GACGATGGCG GAAGGAGAAG
1441   CCTCGAAAAT CAGAAGAGCG ACCAAAGGGA TATTCAAGTA TTCGCCCCTC TTGAAGCTGT
1501   TTATACGGGC GGCTGGGTGT GTGTATGTGT ACTTGAGTAC CTACCTCGTG TCTCCCGTTG
1561   CTATACGATA TGAGCTTCCC CACGACGCGC CTTTATGGCC TGACCAAGGT CTCGATTATC
1621   CGGCTCCTGC GGGTGACACT GCCGAGGGGG GTTACATACG GTCCAGCAGC GGCGATGGAG
1681   TTTGGTCCCT GAAGACTGCA TGGCGGGGCC AAGCGATGAG GAACGCCGTT ACATGCATGT
1741   GCATGTAGAC GCCGCCACCC ATGAGGCCCG GAACAGTCTA TCGAAGCTCA GGGATTGGCC
1801   CGGCAACTCG ACGCCCCGTC GAGCGGCTCA CCGGTAGTCG ACGGCGTCCG TCGGAATCTC
1861   GCGCTGCTGC GGGCCACCAC GGCGATGGGC CGTACACACT GCTACTACGG TGTACAATGT
1921   ATCATGTACC CGATCGACGA GGAACTCGGG GTAGAGGTAC CCCGTACAAT CCAGTTTCTC

```

1981 AACCCTAATGG AACCACACAT ACGGGGTGGC TTTGGTTCAC GTTGCACTTT AAACTCGCAG
2041 ACGAGGGACC GACCTGCAGC GTGGCCCACT TCTGAAGCCT GCCCAGCTTT CTGCAAGACG
2101 CGGGCCATCG CGCTTGGCCG AGGAGAGAAA GGGTATCCAT GGCGACAAAG GCGGTCCTGG
2161 TGGGTTCGGT GCCGGCTTTG GAGTTCACTG GTCTGGGTGG GTGGCCAGCT GGATGCATGC
2221 ATTGGCCTGT ATCAAAGGTC CGGGATTCCC CAGGAGTATA AGACGTTCGT GCTGGGAGAT
2281 CTAGCGACGT GTTGGGAAAT ATCGGCCGTA GAGTGCGAAA AAGAACTGGC GGAAATATTT
2341 CTCCTTGGAC TCGGTCACAC TCAGTCAGTA GTGGACTGCC AGTCTATCAT ACACCTTTGA
2401 TATCAACATG ACTATCCTTA CAGGTGCCGA CGACGCCTCG TCATACCACA GGTATGTCTT
2461 CACAGCCTCT GGAAAGCGCA GTTGGGAGCT ATCTCTAACA TTACCACATC AGGCGCAATG
2521 GAAGCTCTGA TATCCCAAAA GGTGCCATCC ACCGCAACGG CTTCGCAGCC GCAGCCCCTG
2581 ACTGCTGGAT CCGGTCCGTG CTGTTTTCCG TGCACCAGAT GCTCAAGAGG TTCGGAAACG
2641 GATCTCACAC CGTCGTGGCG TCACTCGTAA CTTTCATCAGA GGGATGCCCT TCAACTTCGG
2701 CCTGGAGGGC CATCCCCTCC GTCATCCATC ATATAGAGGG CGGAGACAAC AACAACACAG
2761 TCGCCTCTGC CGTGGAACAG GCGGCGAATC TCCTGAACTC AGAAGGATCG GGACAGGACC
2821 TTCTGATTCC CATCGGACTC ACTGAGCTCG TCAAGTCGGA GCTGATTGAC CTCCTGGTCA
2881 TCTTCGACGA CGAGACAAAT AACATACGAC TGCCGCAGGA CTTCCTCACTT ATCCTGCGGA
2941 TACATCAGCG GCAAGACCAC TGGCAGCTGT CAGTCCGGTA TCCCTCGCCC CTTTTCGACA
3001 CCATGGTCAT CGACAGCTTT CTGAGCGCAC TTCACAACCT GTTGTCCGCG GTGACAAAAC
3061 CCTCCCAGCT CGTGCGCGAC ATCGAGCTGC TCCCAGAATA CCAGGTCGCT CAGCTGGAGA
3121 AGTGGAAACAA CACAGACGGC GACTACCCCA CCGAGAAGCG GCTACATCAT CTGTTTCGAGG
3181 AGGCAGCAGT GCGTCGTCCC CAACACGTTG CCCTCATCTG CGGCGACAAG CGCATCACCT
3241 ATGAGGAGTT GAATGCTATG GCGAATCGCC TGGCCACCA TCTGGTATCC TCGGGTATCC
3301 AGACTGAGCA GCTCGTCGGT CTCTTCCTCG ACAAGACCGA GCTCATGATC GCTACTATTC
3361 TGGGCATCTG GAAATCTGGT GCCGCGCATG TACCTATCGA CCCTGGGTAC CCGGACGAGC
3421 GTGTCAAGTT CGTCCTGAAT GATACGAAGG CGCAAGTGGT CATTGCTAGT CAGAGGCACG
3481 TCGATCGACT GCGGGCTGAG GCTGTTGGCG GCCAGCATCT TCGCATCATC GGTCTCGAAT
3541 CTCTGTTTGA CAACCTTGCT CAACAGACAC AACACTCACC AGAGACGTCG GGCAATTTGA
3601 CCCATCTGCC CCTGAACAGC AAACAGCTTG CGTACGTGAC ATACACCTCG GGCACCACGG
3661 GCTTCCCGAA AGGCATCTAC AAGGAGCACA CAAGCGTCGT TAACAGCATC ACCGATCTGT
3721 CTGCTCGGTA CGGTGTGGCC GGGGAGGACG ACGAGGTGAT ACTCGTCTTC TCCGCCTACG
3781 TCTTCGAGCC ATTCGTGCGC CAGATGCTCA TGGCCCTGAC CACGGGCAAC TCTCTCGCCA
3841 TCATCAGCGA CGAGGACAAG TTCGACCCTG ACACCCTTAT TCCCTTCATC CAAAAACACA
3901 AAGTCACTTA CATCCACGCC ACCTCGTCAG TGTTCAGGA GTACGACTTC GGGTCCTGCC
3961 CCTCGTTGAA ACGCATGATT CTGGTGGGAG AGAACTTGAC AGAGCCGCGC TACGAGGCCC
4021 TGAGGCAGCG CTTCAAGTCG CGCATCCTGA ATGAATATGG CTTACCCGAG TCTGCGTTTG
4081 TGACGGCGCT CAACATATTC GAGCCTACCT CACAGAGGAA GGACATGAGT CTGGGAAGGC
4141 CGGTGCGCAA CGTCAAGTGC TATATCTTGG ATGCCAACCT CAAGAGAGTC CCCATCGGTG
4201 TTACAGGGGA GCTGCACATC GGTGGCTTGG GTATATCCCG GGGGTACATG AATAGGGAGG
4261 AGCTCACAAG GCAGAAGTTC CTCCCGAACC CCTACCAGAC CGATAAGGAG CGCCAACGGG
4321 GTGTCAACTC AACCATGTAC AAGACAGGAG ATCTGGCCCC CTGGCTACCC AGTGGCGAAG
4381 TCGAGTATCT CGGCCGTGCC GACTTCCAGA TCAAGCTGCG CGGCATTCTGA ATTGAGCCCC

4441	GCGAGATCGA	GTCCACTCTC	GCCATGTATC	CCGGAATCAG	GGCCAGCATC	GTCGTGTCAA
4501	AGAAGCTTCT	CAGTCAGGGG	CAGGAGACGA	TCCAAGACCA	CCTTGTGGGG	TACTATGTTT
4561	GCGATGAGGG	CCACATCCCC	GAGGGTGACC	TGCTGAGCTT	CCTGGAGAAG	AAGCTACCTC
4621	GGTACATGGT	CCCAGCGCGC	CTTGTCCAAC	TGGCTCAGAT	TCCAACCAAT	ATCAACGGCA
4681	AGGCGGATCT	GCGTGCTCTT	CCTGCCGTCG	AAGTCGCCGT	AGCTCCCACC	CACAAGCAGG
4741	ATGGCGAGCG	AGGAAACCAG	CTGGAGAGCG	ACCTGGCTGC	CATATGGGGC	AACATTTTGA
4801	GTGTTCCCGC	TCAAGACATT	GGGTCTGAAT	CCAACTTCTT	CCGCCTGGGT	GGCCACAGTA
4861	TTGCATGCAT	CCAGCTCATT	GCTCGTGTGC	GACAGCAGCT	AGGCCAGGGG	ATTACCCTCG
4921	AGGAGGTCTT	CCAGACCAAG	ACGTTGCGAG	CTATGGCTGC	CCTCTTGTCG	GAAAAGTACA
4981	CGAAGGCGTC	GAATGGGACG	AACGGAGTGA	CCAACGGCAC	TGCTCACGTC	AACGGCCACG
5041	CAGCGAACGG	CCATGTCAGC	GACAGCTACG	TGGCCAGCAG	TTTGCAGCAA	GGCTTTGTTT
5101	ACCATTCACT	CAAGAACGAA	CTGTCCGAGG	CGTACACCAT	GCAATCCATG	ATCCACTATG
5161	GTGTGCCCCCT	GAAACGGGAT	ATTTACCAAG	CGGCATGGCA	GAGGGTACAG	GGGGAGCACC
5221	CTGCACTGCG	GCTTCGGTTC	ACATGGGAGG	CCGAAGTGAT	GCAGATCGTG	GACCCGAAAT
5281	CTGAACTCGA	CTGGCGTGTT	GTTGACTGGA	CCGATGTTTC	GAGCCGGGAG	AAGCAGCTGG
5341	TTGCGCTGGA	GCAACTCCAA	ACGGAGGACC	TTGCTAAGGT	CTACCATCTC	GATAAGGGGC
5401	CCCTTATGCG	ACTATACCTC	ATCCTGCTTC	CGGACTCAAA	GTACTCCTGT	CTGTTTCACT
5461	GCCACCATGC	CATTCTCGAT	GGGTGGAGTC	TGCCCCTGCT	CTTCAACAAT	GTCCACCAGG
5521	CCTACCTCGA	TCTCGTCGAA	GGCACTGCTT	CGCCCCGTCG	GCAGGACGCT	ACCTACCTAC
5581	TCGGCCAGCA	GTACCTGCAG	AGCCACAGGG	ACGACCATCT	CGACTTCTGG	GCCGAGCAGA
5641	TCGGCAGGAT	CGAAGAGCGC	TGCGACATGA	ATGCGCTGCT	GAATGAGGCC	AGCCGATACA
5701	AGGTGCCCCCT	GGCCGACTAT	GACCAAGTCC	GCGAGCAGAG	GCAGCAGACC	ATCAGTCTGC
5761	CCTGGAACAA	CTCCATGGAC	GCTGGTGTGC	GGGAAGAACT	CTCCAGTCGT	GGCATCACCC
5821	TTCATTCCAT	TCTACAGACG	GTCTGGCACC	TGGTCCTCCA	CTCTTATGGA	GGAGGCACCC
5881	ACACGATCAC	CGGCACCACC	ATCTCCGGCC	GTCACCTGCC	CGTCCCCGGA	ATTGAGCGCT
5941	CTGTTGGTCT	CTTCATCAAC	ACACTCCCTA	TGATCTTTGA	TCACACCGTC	TGCCAGGATA
6001	TGACAGCGCT	CGAGGCCATT	GAGCATGTCC	AAGGCCAAGT	CAACGCCATG	AACTCCCGGG
6061	GCAACGTCGA	GCTCGGACGC	ATGAGCAAGA	ACGACCTCAA	GCACGGGCTC	TTCGACACCC
6121	TCTTCGTCCT	CGAGAACTAC	CCAAACCTCG	ACACGGAGCA	GCGGGAGAAG	CACGAGGAGA
6181	AGCTCAAGTT	CACCATCAAG	GGTGGCACGG	AGAAGCTCAG	TTACCCGCTG	GCCGTGATTG
6241	CCCAAGAGGA	CGGCGACAGC	GGATGCTCGT	TTACGCTCTG	CTATGCGGGC	GAGCTCTTCA
6301	CGGATGAGTC	CATCCAGGCG	CTCCTGGACA	CTGTCCGGGA	CACCCTGAGT	GATATTCTCG
6361	GGAACATCCA	TGCCCCATATC	CGCAACATGG	AGTACCTCTC	CTCGAACCAG	ACGGCGCAGC
6421	TCGACAAAGTG	GAATGCCACC	GCCTTCGAGT	ACCCCAACAC	CACACTGCAC	GCCATGTTTCG
6481	AGTCCGAGGC	GCAGCAGAAG	CCGGACAAGG	TGGCCGTGGT	GTACGAGGAT	ATCAGGCTGA
6541	CCTACCGCGA	GCTCAACAGC	CGTGCCAATG	CCCTGGCGTT	CTACCTCCTC	TCCCAGGCGG
6601	CTATCCAACC	GAACAAGCTG	GTCGGGCTGA	TCATGGACAA	GAGCGAGCAC	ATGATCACGA
6661	GCATCCTCGC	GGTCTGGAAA	ACGGGTGGAG	CCTACGTCCC	GATCGACCCT	CGATACCCTG
6721	ACCAGCGTAT	CCAGTATATC	CTGGAGGATA	CGGCGGCTCT	CGCAGTCATC	ACGGACAGTC
6781	CTCATATTGA	CCGTCTGCGC	AGCATCACCA	ACAACCGCCT	TCCTGTTATC	CAGTCGGACT
6841	TTGCTCTCCA	ACTCCCGCCC	AGCCCAGTTC	ATCCCGTCTC	AAACTGCAAG	CCAAGCGACC

6901 TCGCCTACAT CATGTACACA TCCGGCACCA CTGGCAACCC CAAGGGTGTC ATGGTGTAGC
6961 ACCACGGTGT AGTGAATCTG TGC GTTTCAC TCTGCCGGCT CTTCCGGCCTT CGGAACACAG
7021 ATGACGAGGT CATCCTCTCG TTCTCGAACT ACGTCTTCGA CCAC TTTGTC GAGCAGATGA
7081 CGGATGCCCT TCTCAACGGT CAGACTCTTG TGGTCC TCA CGACGAGATG CGTGGCGACA
7141 AGGAGAGGCT TTACAGATAC ATCGAGACCA ACCGCGTCAC GTACCTCTCG GGGACACCTT
7201 CCGTCATCTC CATGTACGAG TTCGACCGGT TCCGCGACCA CCTGCGGCGC GTGGATTGCG
7261 TCGGCGAGGC CTTCAGCGAG CCGGTATTCG ACAAGATCCG CGAGACGTTT CCGGGTCTCA
7321 TCATCAACGG TTATGGCCCG ACTGAGGTGT CTATCACTAC CCACAAGCGC CCCTACCCGT
7381 TCCCGGAGCG CCGCACAGAC AAGAGCATCG GTTGCCAGCT GGACAACAGC ACGAGCTACG
7441 TCCTCAACGA TGACATGAAG CGCGTGCCCA TCGGGGCCGT GGGAGAGCTG TACCTTGGTG
7501 GCGATGGCGT CGCTCGCGGA TACCACAACC GGCCAGACCT GACGGCTGAC CGGTTCCCTG
7561 CCAACCCCTT CCAGACGGAG CAGGAGAGAC TTGAGGGCCG AAATGCGCGT CTGTATAAGA
7621 CTGGTGACTT GGTTCTGCTGG ATCCACAATG CAAACGGCGA TGGTGAGATC GAGTACCTCG
7681 GCCGCAACGA CTTCCAGGTC AAGATTCGAG GCCAGAGAAT CGAGCTGGGA GAGATCGAGG
7741 CCGTGCTTTC ATCCTATCCG GGCATCAAAC AATCCGTCGT CCTGGCCAAG GACCGCAAGA
7801 ATGACGGGCA GAAGTACCTC GTCGGCTACT TCGTCTCCTC AGCAGGGTCC CTGTCCGCCC
7861 AGGCCATCCG CCGCTTCATG CTCACGAGCC TGCCCGATTA CATGGTTCCT GCGCAGCTGG
7921 TGCCCATCGC CAAGTTCCCC GTCACCGTGA GCGGGAAGCT CGATGCCAAG GCCTTGCCCG
7981 TGCCAGACGA TACAGTCGAG GATGACATTG TGCCACCGCG TACCGAGGTT GAGCGCATCC
8041 TAGCTGGGAT CTGGTCTGAG CTGTTGGAGA TACCGGTCGA CAGGATCAGC ATCTACAGTG
8101 ACTTCTTCAG TCTGGGCGGC GACAGTCTCA AGAGTACCAA GCTGTCCCTT GCTGCCACGC
8161 GGGCTCTCGG TGTGGCCGTC AGTGTCCGCA ACTTGTTTCA CCATCCGACT ATCGAAGCCT
8221 TGTCTCAGTG GATTATCAGG GGTTCGAACG AGGTCAAGGA TGTGGCTGTG GTGAAGGGCG
8281 GTGCCAGTCT TGATATCCCC CTATCCCCTG CCCAGGAAAG ACTCATGTTC ATCCACGAGT
8341 TCGGCCATAG CGGCGAGGAT ACTGGTGCTT ACAATGTGCC TTTGCAGCTG CAGCTTCACC
8401 ATGATGTCTG TCTCGAGTCG CTTGAGAAGG CTCTGCGGGA TGTCGTCTCG AGACACGAGG
8461 CTCTCCGGAC CTTGATCACC AGGACCCAGA AGTCCTCCGT GCACTGCCAG AAGATCCTCG
8521 ACGCCGAAGA AGCGCAAAAG CTCTTCTCTG TTGATGTTCT GCGCCTGACC TCGGAGACGG
8581 AGATGCAGGG CAGGATGGCC GAGAGTACCG CCCACGCCTT CAAGCTCGAC GAGGAAC TCC
8641 CGATTCATGT ACGCCTGTAC CAGGTTGTAC GTGATGGCCG CACGCTCAGC TTTGCCAGCA
8701 TCGTCTGCCA CCATCTGGCG TTTGACGCGT GGT CATGGGA TGTGTTCCAG AGGGACTTGG
8761 ACGCCTTCTA TGCCGTCCAT ACGAAGCACA AGGCTGCCGC CAACCTGCCA ACCCTCCGCG
8821 TGCAATATAA GGAGTATGCG ATAGAGCACC GCCGGGCTCT CCGCGCTGAG CAACACCGTG
8881 TTCTCGCGGA CTA CTGGCTG CGCAAGCTCA GTGACATGGA GGCGTCTTAT CTGGTCCCCG
8941 ATCGCCCTCG ACCGGCGCAG TTTGACTATA CCGGGAACGA TCTCCAGTTC TCAACTACTC
9001 CCGAGACCAG CGCGCAGTTG AAGGAGCTGG CCAAGCGCGA GGGTTCAAGC CTCTACACCG
9061 TTGTGGCGGC GGCGTACTTT CTGCTTCTCT ACGTGTACAC CAACCAGCGG GATATCACGA
9121 TTGGTATTCC CGTTGCGCAC CGTAACCATC CGGACTTTGA GTCGGTTGTC GGCTTCTTTG
9181 TCAACTTGCT CCCTCTGCGG GTCAACGTGT CTCAGTCGGA CATTCATGGA CTTATCCAGG
9241 CAGTGCAGAA AGAGCTTGTC GATGCCCAGA TCCATCAGGA CTTGCCATTC CAGGAGATCA
9301 CCAAGCTTCT TCATGTGCAG CACGATCCAA GCCGCCATCC CCTTCTCCAG GCCGTGTTCA

9361 ACTGGGAAAA CGTACCCGCC AATGTCCACG AGGAGCAGCT GCTTCAGGAG TACAAGCCGC
9421 CCTCGCCTCT GCCTTCGGCG GCCAAGTTTG ATCTCAACGT CACGGTGAAA GAGAGCGTCA
9481 ATTCGCTCAA CGTCAACTTC AACTATCCTA CCAGCCTCTT CGAGGAGGAG ACCGTTTCAGG
9541 GGTTCATGGA AACCTTCCAT CTCCTTCTTC GACAACTGGC CCACAACAAG GCTAGCACAA
9601 GCCTCTCGAA GCTGTCGGTT GAAGATGGAG TGTGAATCC AGAGCCGACT AACCTTCAGC
9661 CCTCAAGCCG GGACAGCGGA AATTCACTCC ATGGGCTCTT CGAGGACATC GTGGCCTCGA
9721 CCCC GGACCG CATCGCAATT GCTGACGGCA CCAGGAGTCT CTCGTACTCC GAACTCAACG
9781 AGCGGGGCAA CCAGCTGGTA CATTTGATCA TCTCTTCTGC CAGTATTGTA GCAGACGACC
9841 GCATCGCTCT TCTTTTGGAC AAGAGCATCG ATATGGTGAT TGCTCTCCTG GCAGTTTGA
9901 AGGCCGGTGC CGCATATGTG CCCCTTGACC CGACATATCC GTCGCAGAGG ACTGAGCTCA
9961 TCTTGGAGGA ATCTAGTGCC AGGACGCTCA TCACCACTAG AAAGCACACG CCGAGGGGAG
10021 GAACAGTCGC AAATGTTCCA TCCGTGGTCC TTGACAGCCC CGAGACCCTA GCCTGCCTCA
10081 ACCAGCAGTC AAAGGAAAAC CCGACAACGT CAACGCAGAA ACCGTCCGAC CTCGCATATG
10141 TCATCTTCAC CTCGGGAACC ACAGGCAAGC CCAAGGGGGT TCTGGTGGAG CACCAGAGCG
10201 TAGTCCAGCT GCGCAATTCC CTCATCGAGC GATACTTCGG CGAGACCAAC GGGTCTCAGC
10261 CCGTGCTCTT CCTGTCCAAC TACGTCTTCG ACTTCTCTCT TGAACAGCTC TGTCTCTCAG
10321 TCTTGGGTGG AAACAAGCTC ATCATTCCAC CAGAGGAGGG TCTCACGCAC GAGGCATTCT
10381 ACGACATCGG CCGCAGGGAG AAGCTATCCT ATCTCAGCGG GACGCCCTCG GTGCTGCAGC
10441 AGATTGAGCT CTCCCGTCTG CCGCATCTTC ACATGGTCAC CGCTGCGGGC GAGGAGTTCC
10501 ACGCTAGTCA GTTTGAGAAG ATGCGCTCCC AGTTCGCGGG CCAGATCAAC AACGCCTATG
10561 GTATCACTGA GACGACCGTG TACAACATCA TCACCACGTT CAAGGGCGAT GCCCCTTTA
10621 CCAAGGCACT CTGCCACGGG ATCCCCGGAA GTCACGTCTA CGTCCTGAAC GACCGACTTC
10681 AGCGTGTTCC TTTCAACGCT GTTGGCGAGC TCTACTTGGG CGGTGACTGC CTTGCTCGCG
10741 GGTACCTCAA CCAGGATGCC CTGACCAACG AGCGATTCTT CCCCACCCTT TTCTACGAGC
10801 CGAAACAGGC AAGTGACAGT CGTCCCCAGA GACTCTACAA GACTGGAGAT CTGGTGCGCT
10861 TCCGTGGACC CCACCATCTC GAGTATCTCG GCCGCAAGGA CCAGCAGGTC AAGCTGAGGG
10921 GCTTCCGCAT CGAGCTCTCC GAGGTGCGGG ATGCCGTCTT AGCCATCTCT GCTGTTAAGG
10981 AGGCTGCCGT CATCCCCAAG TATGACGAGG ATGGCTCCGA TTCACGAAGG GTCAGCGCCA
11041 TCGTCTGCTA CTACACGCTC AACGCCGGAA CTGTGTGCGA AGCATCGAGT ATCCGTGACC
11101 ACCTGCACGC CAACCTTCCC CCGTACATGG TCCCAAGTCA GATCCACCAG TTGGAGGGAT
11161 CTCTCCCCGT GACCGTCAAT GGAAGCTCG ACCTGAACAG GCTCTCCACA ACTCAAGTCT
11221 CGCAGCCAGA GCTTTACACC GCTCCACGAA ATTCGACAGA GGAAACCTTG TGCCAGCTTT
11281 GGGCATCTCT CTTAGGCGTC GACCACTGCG GCATTGACGA CGACCTGTTT GCCCGAGGCG
11341 GCGACAGCAT CTCCTCTCTC CGACTAGTGG GTGACATCTA CCGCGCGCTA GGACGCAAGG
11401 TCACCGTCAA GGACATCTAC CTCCACCGCA GCGTCCGAGC CCTAAGCGAA AATGTCCTGA
11461 CCGACCAGAA GGATAAGGGT ACTCTGCCAG CGTCTCCTCC CCTCCAGCGA GCGGAGCAGG
11521 GCCAGGTTGA GGGCGACGCA CCGCTTCTCC CCATCCAGGA CTGGTTCTTT TCCAAGCCCC
11581 TGGATAACCC CGCTTACTGG AACCCTGCTC TCACCATTCG AACCGGGGCA CTCTCCGTGC
11641 AAGGGCTCCG GGGTGCTCTG AAGCTGCTGC AGGAGCGCCA GCACGTGCTG CGTCTGAGAC
11701 TGCAACGCCG GGACGAAGGT CGCCATGTTT AGACCTTTGC GCGTGACTGC GCGCAACCTC
11761 GCTTGACTGT GCTAGACCGA CGAAGCTTCG AGGACGCAGA GGATGTACAG GAGGCTCTCT

11821 GCGAGATCCA ATCTCATTTT GACCTCGAGA ATGGACCCCT CTACACAGTG GCGTACATCC
11881 ACGGTTACGA GGACGGCTCC GCCCGAGTGT GGTTCGCTG CCATCACGTC ATGGTCGACA
11941 CTGTGAGCTG GAACATTATA CTGCAAGACC TGCAGGCTCT CTATCATGGA GACAGCCTTG
12001 GTCCCAAGAG CAGCAGCGTG CAGCAGTGGT CGCTAGCTGT CAGCGACTAC AAAATGCCAC
12061 TGTGCGGAGAG GCGGCATTGG AATGTGCTCA GGAAGACAGT CGCCCAGAGC TTCGAGACCC
12121 TGCCTATCTG CATGGGCGGC GTGCTCCAGT GCCAGGAGAA GTTCTCGAGG GAAACGACAA
12181 CAGCTCTGCT CTCCAAGGCC TGCCCTGCCT TGGACTCCGG TATGCATGAG ATCCTTCTCA
12241 TGGCCGTGGG CTCCGCGCTG CAGAAGGCGG CAGGGGATGT CCCTCAGGTC GTCACGATAG
12301 AGGGTCACGG GCGCGAAGAT ACTATCGACG CAACTCTGGA CGTCAGCCGG ACAGTCGGCT
12361 GGTTCACGAG CATGTACCCC TTCGAGATCC CCAAAGTGAC CGACCCCGCT CAGGGCGTCCG
12421 TCGATGTCAA GGAGGCGATG CGTCGCGTGC CGAATAGGGG TGTCGGTTAC GGTCCAGCCT
12481 ACGGATACGG CGGATCGTCG CTGCCC GCGG TGAGCTTCAA CTACCTTGGT CGCCTGGACC
12541 AGGCTTCCTC GGGGGCTCAA AGGGACTGGA CGCTGGTCAT GGATGAAGAC GAGTATCCGG
12601 TCGGACTGTG CACCAGCGCT GAGGACTCGG GACGAAGCTC CTCCATGGTG GATTTACCT
12661 TCTCTATCTC TGGCGGCCAG CTTGTCATGG ATATGAGTAG CAGCTGGGGC CACGGCGCAG
12721 CAAATGAATT CGTTCGCACA GTTCGTAACA CACTAGATGA CTTGATCAAA ACAACGAGCA
12781 GCAGGGACTT CAGCGCACCT CTGCCCTCCGT CGGATCAGGA GTCCAGCTTC ACCCCTTATT
12841 TTGTCTTCGA AGAGGGCGAG CGACACGGCG CTCCGCTCTT CCTGCTCCCA CCTGGCGAAG
12901 GCGGAGCGGA GAGCTACTTC CACAACATTG TCAAGGGTCT CCCGAACCGC AATCTTGTCTG
12961 TGTTCACAAA TCATTACCGC GAGGAGAAGA CGCTCCGGAC CATCGAGGCG CTGGCCGAGT
13021 ACTACCTGTC GCACATCCGA TCCATCCAGC CGGAGGGGCC ATACCACATC CTCGGCTGGA
13081 GTTTCGGAGG CATCCTCGGT CTCGAGGCGG CAAAGCGATT GACTGGCGAG GGTACACAAGA
13141 TTGCCACGCT GGCACCTATC GATCCGTAAT TTGACATCCC GTCCGCGTCC AAGGCCATCG
13201 GCCAACCTGA CGATGCCTGC GTCTTGACC CCATATACCA CGTCTACCAC CCGTCGCCCGG
13261 AGAGCTTCAG GACGGTGTCA TCTCTCACTA ATCACATAGC CCTGTTCAAG GCTACCGAGA
13321 CGAATGACCA GCATGGCAAT GCCACGCAGC AGGCCCTGTA TGAGTGGTTT GCCACGTGCC
13381 CTTTGAACAA CCTGGACAAG TTTTGGCGG CCGACACGAT CAAGGTGGTT CCTCTGGAGG
13441 GTACACATTT TACCTGGGTG CACCACCCGG AGCAGGTGCG CTCAATGTGC ACTATGCTGG
13501 ATGAATGGCT TGGGTGAACG AGGCAGTTGC TGTGAGAGAA TGAGAATGAG ACACAAAACG
13561 CGGGCGGAAG AGAGACTTCC TCGGACGGCG GGTTCGCC GACGAGTGAT GACCTGGTCC
13621 CAGGGGTCTG GTGATATTTT CTCCTGAATG TGTGAGGATA TTAGTGGTTT TTTTCTGCCG
13681 TTAGAGACGT ATTTAGTAAG CTCTGGAGTT TGGAGTCATT ATTTCTCTGA ATGGTCTTCT
13741 TCTGTAGTAA TAAACTAGCA GAGCGGATTA TATATATATA TATATATATG TATTTGGCTG
13801 GTATTGCTAT GCGTGTTCTT ATGTGAATTG GTATATGTAT AAGTATGTCT ACCTTACTGC
13861 ATCTGTTAAT TCTTATGTAC TGCTACATGA GTTGCTACGG TTATTGCGAC GTGATGCGTG
13921 TAGTCGAAGT TATTGTACCT ATTGCGTGTA CTATGCTCCC CTCTTCTTTC TACTATATCT
13981 CGGTGTAGTA CAAACAAAAC GACGTAGGGG AAGTGGGGAA GAAGTTGAAC GAGTATAGAC
14041 TCCCCGAGCA ATGAATCAGT ACATTATATA TACTGTTTCT TCTCTCTCAT GACTTGGTAC
14101 GTGGGAAGTT CCCAACATCT AACCTGTCCA ACCCTCCGAC CAGAAAGCTG TCCAACAATG
14161 TCGTCCAACG CTCCCAGCTC CCCAAGTATC ATCCATGGCA CCGAAGGAGC CGCATCATCA
14221 ACTGAAAGGC ATCATCTTTG GACTATCGCA CCGGGATAAT TCTATACGAC ATCCAAGTTG

14281 AATTGTCTGA GCCTCAGGTC AGGATCCCAG GCGTCAATCT TGGCAGGGGA TAGTGAAACC
14341 TTGAAACAAG GCATGCGGGC ATTGCGGGCA TTAGGCTCGC ATAGAGAGGG GGGGTTCAG
14401 ACCGAGAGAT ATGCGTGCAG GGCCACGATA TTGGTGCGCC CAATTACTAG CATAAAATAC
14461 TAATACAAAT GCAGGCAACC GAAGGGGATT ATTCATAAAT GCCTTGGGTG CATGGCAACT
14521 CGAAAAGTTG GGAATGTGG CAGGAGCGCC CAAGCATATT CACTCGTCAA GAGTCGTCAT
14581 CACCACCCCC GCCCACCTTC CTATCTTATT TCCTCCTCCT TCAGATCGTA CGACACTCAT
14641 CTTCCCTCGT CTCTTTTCCA TCGTCGTCAT CAACAAAAGT CATCTGGCTC GTCTCCATCC
14701 GAAACACCTA CTAGTAGCTA TTAGTCACTC ACTATGAGAC TCTCCTTCTT CGCGCCAAAC
14761 CGTACGATAA AAAGTCCTTT GGCCTCGCCC ACGCCGCCAG GGATCCGCCG TCACCCGTCA
14821 CGATAACCTA CCACGACATC CCCCTCAATG AGGACACCGT CTCCACGGTT AGGGACGCCG
14881 ATGCCGTCTG CGCCTTCGTA AATGACTCCC TCTCCGCTCA CGTCATCGAG ACCCTCGCCA
14941 GGCAGGGTGT CAAGGCCATC CTCCTCCGCT GCGCCGGCTT CAATCACGTC GACCTCGCCG
15001 CCGCCGCCCG ACACGGCATC ATGGTCGCCA ACGTGCCGTC GTACTCGCCA GAGGCCGTCG
15061 CCGAGTTTCG GGTAGCCCTG ATCCAGACAC TCAACCGCAA CACCCACCGC GCCTACAACC
15121 GCGTGCGCGA GGGCAACTTT GCCCTCCACG GCCTCCTGGG TAAGACACTG CACGGCAAGA
15181 CGGTGCGGGT AGTGGGCGTG GGCAAGATCG GCTTGGCCAC GGCGAGGATC ATGAAGGGCT
15241 TCGGGTGCCG CGTCCTTGCT AGCGACCCAT TTCCCTCGCC TCGGTTTGAG GAGTACGGCG
15301 AGTACAAGGA CCTGGACACG TTGCTGTCCG AGTCGGATAT TGTCAGCCTC CACTGTCCCC
15361 TCATGGACAA CACGCGGCAC ATCATCAACG GCGACACAAT CGCCAAGATG AAAAAGGGTG
15421 TCCTCCTCAT CAACACGTCC CGCGGTGGCC TTGTGGACAC CCGTGCAGTC ATCAAGGCCC
15481 TCAAACAAA GCACATTGGC GGCGTAGCCC TCGACGTGTA CGAGGCCGAA GGCTCACTGT
15541 TCTACGACGA CCACTCCGGT GAGATTATCC ACGACGATGT CCTCATGCGC CTCATGACAT
15601 TCCCCAACGT CATCGTCACG GGACACCAGG CCTTCTTCAC CGAAGAGGCG CTCGAGGAGA
15661 TTGCACAGTG CACGCTGCGT AACCTGGAGG AGTTCATCAA GGAGGGAACA TGCAAGAACT
15721 CGCTGACCAA GGAGCCCGAA TTGAGGTCCA AGGTCCCTGA CCCGGTGCGC AATGTTTAAA
15781 TTGATGTGGA TGATTGAATT CTATATTCTG GTATCTCTGT CTATGTACGG TCATCTGAAA
15841 CTTTTGATGC TGGTTAAATG GTGAGTCCTG CTAACGCCAC ACACAAACAC ACACACGCAC
15901 ACACACAGNG CATATTGAGA CGAAACTGGG GAAAGCTAAG TATCAATAAA ACACAAAACG
15961 AAATGGACGG AAGGATATCT CCCGCTCTAG TATATAAGGC GTACGAAAAC ACCCGTTGTA
16021 CAACCGCTTA AG

Figure 6 (SEQ ID NO 6): A. chrysogenum genomic DNA sequence of an approx. 5.8 kb region marked by EcoRV and BamHI and containing the biosynthesis genes *cefD1* (position 2372 to position 180, inverse arrangement) and *cefD2* (position 3888 to position 5133). A single strand in 5'-to-3' orientation is shown. The particular translation start codons (ATG) and the particular translation stop codons (TAG) of the respective coding regions are depicted underlined and in bold type. The intron *cefD2* and the said cleavage sites are depicted underlined.

```

1      GATATCTGAG TGTTGTTCC GCGCGTTACG GAGTCATAGC CGGTTGAGTT CGGCCGAATC
61     TGCCCTCTAT GTTGTGTTTA CGCTCTTCAG GCTTTATGCG GAATTACATG TGTTTTGCAA
121    GCCACATTTG TTTTCTATGT TACATCGTTT CACGCGTCGA TGGGTCCCCT TAATAGCAAC
181    TAGAGTCTTG GCCGAGCCGC CGACTGGCCG GACAGTAGAT CCCAGTCCCCT CTCGGTAAAG
241    GGAACATACT TGCTCGCACC GGCCGGCAGC CAGAAGAAGC GATCCTTTTC ACTTCCAGGG
301    ACCTTGCTGC CCATGGAGCG CGGGTCGACC CCCTCGTCGC GGAGTGGCAC CTTGTTGTGT
361    TTATGGTTAT CCGTGCTCAT GCCTCCACC GTCTCCCGCA CCCTAATGAA GACAGGAACG
421    GCGTAGGAGG GCAACTCTGA CCGCAGGAGC GACGTCAGCC GTGACCAGTC CAACGTGTCT
481    GGGGTTGCCG CGTTCTTCAA TGCGATGGCT GCGCAGCCGG CTCGCCCCTC GTGGTTGGGG
541    ACCTGGACGC CGTAGACATT GGCCTCGGCA ATGTCTGCGT GTGAGCCCAG GACTTGTCTT
601    ACCTCGGTCG TGGACACGTT TTCTCCTTTC CAGCGGTATG TGTGCGCTGG TATTCACGTT
661    TAGCATTGTC ACTCTGCGAT TGTAGCAGGG ATAACGTACC AAGTCGGTCG AGAAAGTACC
721    AGTGACCGTC GGCATCGCGT CGAAGTACCA GTGACCGTCG GCATCGCGTC GAAGAGCATC
781    ACCCGTCCGG AAGTACAGAT CCCCCTTCTC AAACACATTC TCCACCAGCT TCTTCTGCGT
841    CGCCTCCTCA GCATGCCAGT ACCCAGCCCA GGCTGACCGT GACGGGAGCC GCGCCAGGAT
901    CTCACCGCCC CTCTCGTACG GCAGTCGCTC GGCAAAGCCC GTCTTCGGCG ACCGCCAGAT
961    GTCACCCGTC TCCGGGTCGA TCCTGACGGG TACGTAGTCA TTGTGAAACT TCGCGCGGAG
1021   GAGCCAAACCG TGGTGGCCGA CGGCGCCGAG GCCGAAACCA CCGCCGCGGT AGTGTTTTAA
1081   GAGTGTCAGG ACGCCTTCTG TGCTGGCGTA GAATTCACCG ATGTCCGATA CGCCGAAGCG
1141   GTCTTGTTTG TCGCTTGGTG TCAATCCTAT GGTTCGATG GGGCTGGACC GAAGTAAAGG
1201   GGGATCATAC CTGGAATTTG GTCCAGAGTT CCGGGCTGAG TCCGTTCCCC CAGACGAGAC
1261   GGACGCGGTG CTGACGGTCC TTTGGTGAGG CCGGAGCAGA GAGTAGGTAC CGAATAAGTT
1321   CCCCRACTGT ACACTCGTCA GGAGGGCAAA CAAGTCGAAG AAGGGTACCA AGGGAGAGTA
1381   GTGGCTTACC ATAGACAAAT ATTGTTGACC CGCTCTCTAT GCAATCGTCC CAGAAGCGAG
1441   ACAAGGAGAA CTTGGGCGCA AGAGCAATGG ATATTCCGCT CATCAAGTCG TTCATGGCCG
1501   CGATGCCCCC CGTTCCGTGG TAGAGCGGGA TGCAGTAGTA GGTGCGGTCA CCGTTGGGGC
1561   CCGGTTTCTG CCCAAATGTC TTTGGCAGCA GTGAAGCAGA GGGATAGTTC CTGGCCACTG
1621   TGATGGGTGC GGCCTTTGGC AGGCCAGTCG TACCACTTTT GCCGTGAGAA TGTGCAACTG
1681   GTGGGTGAAG TCCGAGGCTT TACTCACCTA GTGTACATCA AAGCAAACGG TAGCAGCACC
1741   TTGGTGTCTT CGAAGCAATC TACCGGTGCC CTGTGTGTTT CTTTCCGGGC AATGTCTTCC
1801   TTCAGCGTGC CGGAGAGCAT GATAGCCTCG ACATTGATAT CCCGAAGCCG CTCACCAACT

```

1861 TCGTGAATGC GAGAGGAACA ATCGGACGCA TCATCGTATA TCAGAAACCG CGAGCGCGAC
 1921 AAGCGGACAC AGTGAACCAG AGCATCAGAG CCGAGATTGT AATTGATGAG GGCAGGAGCA
 1981 GCACCGATAG ACAGTAGACC CATCCAGATG AACATCAGCT CCGGCGAGTT GTACAGATAA
 2041 ACACCGACGT GCTGCCCCGC CACCACGCC AGATCGCGGA AGTAGTGGCC ATACTGACAT
 2101 GCGCGTTGGT ACGTCTGCGT CCACGAGTAC TCGGGGTGCC CGCGCGACCA GATGCAAGGT
 2161 GCATCGCCGA GCGGGGAGC GCGGGCCTCG AAGAGGAAGA AGCCAGATGC CTTGCGCTGC
 2221 TCAACAGCTC TGGCGAAGTT CTGGGCCCC CGTTCGGCGC GGGCGAGCTG GTTGAGGTCC
 2281 TTGGTGAGGT GAAGCTTGGC GTCGAGGTAG GCAGCGGCGG CGGTGCTGGC GGCAGCGGCG
 2341 CCAGCGAGGG TTAGTAGGCC GCCGGGTGCC ATTGGGGATT GCCGTTATGC TCCCGTGTGT
 2401 GCTTGTGTCT TTGCGAGCGA TGTCTTGCC TCCGAGACAT TGGTGCTGGT GCTGGTGGCT
 2461 TTGACTGTCT CGGAGAAGGG TGTAAGTGAT CGTGCGATCG GGAGTCCAAA AGTTGACCAA
 2521 ACACGGACAC AACCAGTATT GACGAATAGT TGTTGAAGAC TGCTAGCCTC GGCACCCAGA
 2581 AGCTTGGGAG ACCTGATGGT GTGGCGGGTC TGTGGGTTTG GGCCCTTTTG GGAGGGGGCC
 2641 AAGGGACATG GATGGAGGGA GATCGGCGGC GATGTTGCGG CCGCACTAAC AAGAGGTGTC
 2701 CGTACGTCCG GACTCCGTAC GGTACTGTAC GCGGCCCGTG GAACCAGGCA CTGAAGATTC
 2761 AAGGACTGTT CCTGTCGGAT ATGGCCGCGG CCATGCGTGT CCAGGTTATA GTCTCGACTA
 2821 CATACAAACA TTGTACCTTA CGTAAAAGAA CATGGACAGT AATACGCTGA CTCTTGGCTA
 2881 CGGGATGATA TTTACTGCCC AAGACCGACG CCAAGCAGCG CATACTGTAC GCATAGGCTA
 2941 CGTGGTGGA ACACCGTGGC CAACTCATCC CTTGCTACGG TACAACAGCC GGACAGGGGG
 3001 TGAGGGCGGG TCGGCTCCGC GGGCAGAAGT GCTTACAGCT TAAGAGCTAG TGGTTAAGGT
 3061 TACATGTACT AAGCATTGGT CTGGATCGAA TCAGATGCTG TGCCTTGACT GGATCGAGCG
 3121 GGCCGGCCCC CTGCGTCTTA ATATAGCAAG TACCCGTGAC TATGTAAGTT GTACATGTAC
 3181 ATCTGTGCTG CCACAGGAGC GCCTATAATG TACACCCACC TTGCACGAAC ATTACATTGA
 3241 TACTTGCGTT TCTCCGTACA TGACAGGGGG GTGATATTAC AGTACATGCT CCGCCAAGTA
 3301 ATACAAGACA CGGACCATCG GAGGCAACA TTTGTACTGC AGAATGATGC CTGATTTAGG
 3361 CGCACAGTCT GCATACATGC CATGCCATAA TGAGGTGCTG TATCCGTAGT CTGTGCTCTG
 3421 GATGACTATG AATTATGTCT GAAGTTATTA CTTGGCAACA ACCTGCTTGC CCAACTGGGC
 3481 AGGTTTCATT CATGATGGGA GGATGGAGAC GATGAGTCTA TCCTTGGACA CTGGCATGCG
 3541 CCCCCTCCTT GGTGCAGCTA GATTTTCAGC TTCGATGCAC AGGCTCCGCC CTTGATACTA
 3601 CAGCAGAGTA CATGCCGAGA CAAGAATCTT CAACGTCCCC GATGCGCTTT TGATATCCAC
 3661 CCATGTTTTT ATAGTCGCTG CGCGGTATCG GACCCGAGTC GTCTGTACAC GTGGGTAGAA
 3721 GTTACGCGAA TGCCACCTCA CCACGCTGCA CCATCCATGC CGAGGCAGGC TTCGAGATTG
 3781 CAAGTACGGA CTACAATAAC ATCAAATGGC ATGGTGGGAT TTCCGCTGAC CACCTGCCGA
 3841 CATTACATGT TGTAGTCTTG ACCATTACG CGGTAAATCC CACCTCGATG GACCCCTCTC
 3901 GCCCACATCC GCTATCCGGC AAGCTTGTCG TAGAGCTCGC GGGGCTAGCC CCAGGTCCGT
 3961 CATCCATAGC CACTTTCATG TTGCTCTGCC GAGAGGCCTA CGCATAGAAG CCGCTTCTTG
 4021 CGCTCACGCA CTAACTCGCG CACTCCAGGC CCATTCTGTG GCATGCTCTT GGCAGACTAT
 4081 GCGCCTCAG TACTCCGCAT CGACGGACCG CGATCCCCAA AGGGGGACGT CCTGGCGAGG
 4141 AACAACTCGT CCATCTGCAT CGACTTGAAG CATCCGCCCT CACGCAAGGT GCTCCTCTCC
 4201 ATCCTGTCCC GCGCGGACGT GCTCATCGAC CCGTTCCGGC CCGGCGTCCT GGAGCGTCTG
 4261 GGGCTCTCCC CCACAGAGGT CCTTCTCAAG GCGAATGCCC GCCTGGTGGT CGCCCGTCTC

4321 ACCGGCTTCC GCCGAGATGG CAAGTACCAG GACATGGCAG GCCATGATAT CAACTACCTC
4381 GCCGTGTCTG GCGTCCTGGC TATGCTTGGT AGGGCAGGCG AGAATCCCTT CCCGCCGGCC
4441 AACATCCCTCG GCGACTTTGC CGGAGGGGGC GCCATGTGCG TCGTGGGAAT TCTGCTGGCG
4501 CTCGTATCGC GCGATGCCAC GGGGCTTGGC CAGGTCGTCTG AGGCCAACAT GGTGGACGGG
4561 TCTGCGTACC TGGCCACGAT GCCGCGCCTG GCGACCAAGA CGCCCTTCTG GGGTTCCCCG
4621 CGGGGCGAGA ATGTCCTGGA CGGAGGGTGC CCCTGGTATG CGACATACCG GACAAAGGAC
4681 CCCGGCGGGA AGTACATGGC CGTGGGAGCG CTGGAGCCTC ACTTCTACGA GGTGCTGGTG
4741 CGAGGTCTGG GCCTGGACAA GACGGACCTG CCTCCGCGGG AGGATAGGGC CAATTGGCCG
4801 AGACTGAGGG CGCTATTCTGA GGCAAAATTT GCGGAGAGGA CGCGCAGCGA GTGGGCGGAG
4861 GTCTTTGACG GGACGGATGC CTGCGTCACC CCGGTCCTGG AGCAAGGTGA GCTGGAGAAG
4921 GCCGGCTTCG AACAACGGCT TCCCCTGAAT TTGGGGGCCA CGCCGGGAAA GCCTATTCTT
4981 CCCGGACAGG GTGATTGGAC GGGGCGGACC CTTGCCAAGG GCCATGGAGG AGAGGAGATC
5041 CTGCGCCGGT GGATTGGGTG GGAAAGGGGG GTTGACTACC ACGTTGAGGA GAATAGCGGA
5101 ATTCTCGTCG CTTGTTTCGCG GGAAAAGTTG TAGGCAGGCA GGCAGGCCAT GCTGGTACAT
5161 GCATGCAATG TTGGCCGCTT ATGTACGTAT GTGCATACAT AAACATTGAC AATAGTGGTG
5221 TCATGAAGGA GGAGGGGGGG GGTGGGTTTC GGCCCCTGAC GGTGGCTTGA TCGGGACATG
5281 GACGCCGCAT CGTCAGCGGA GTCAAGCCTC CCGACAGACC TGCCGACCCG ACATCCGAGT
5341 ATCTGTACGT AAGATTGCAT ACCAACAATG TACACCTACT TCCTACGGTT CCAGGATTTT
5401 CTCTCTGGAG GTTGCATGGA GCGAAATGAA AAAAGAAATG CTGCTGCCTG TGTGGAATT
5461 CACAGCCTTG GGATGCGACC TCCCTTTTCA CCGTCTTCAT CCTCGTTGGG GACTCTCAAC
5521 CATGCTGCTG TCGCACTGAG ATAATACAAG GCGTAATTAC TGCAGCGGTA CGTCGTAATT
5581 GGACTTACTT TTGTACGACA GTTGATGTCC AGCGGCATAA AAAGCCTCAG CCGCCAAGAC
5641 TGGCAGACTT CTGCAGCCCT ATCTTGATAT GATCACCCCA TAGGCCGAGG CCCTGTGCTC
5701 GAATCCCGCA CGAAGCCGGA TTCATGTGTA TTCCCAAGGG GGTGAGGACG GAACTCTTAT
5761 TTCGACCTCC GGGGGGCCGA GTTCTAGTCC GCTAACCTTC ACGGCTACAC CGTCCCTCGC
5821 GTCTCAACTA GCCATATAAG TCCTAGGTAA AGAGGTTAAA GTAGGTAGGA AAGGAACCTG
5881 TGGCTTGGCG GATCC

Figure 7 (SEQ ID NO 7): *A. chrysogenum* genomic DNA sequence of an approx. 4.6 kb region marked by XbaI and SgrAI and containing the biosynthesis genes *cefEF* (position 1118 to position 122, inverse arrangement) and *cefG* (position 2058 to position 3534). A single strand in 5'-to-3' orientation is shown. The particular translation start codons (ATG) and the particular translation stop codons (TAG, TGA) of the respective coding regions are depicted underlined and in bold type. The two introns in *cefG* and the said cleavage sites are depicted underlined.

```

1      TCTAGATACC TTGAAACTTG AAACCTCTGA CACTCATGCT TTGTGTTTTA AAAATTTAGC
61     AATTTATAGG GCAATTTTTC TTCTTAAACT CCCGTAGATT TATTACTCGA TCGGCGGTTT
121    CCTAAGTGGC TATAGGAGCT GCGGTAGAGA CAGGGGCAGC CGCGGGGACA GCCGCCTCCG
181    CTGCCGCCGG CTTATCCCTC CGCATGTTGA CATAGTTCCC GCCAAGCCAC TCCCTGAACG
241    TCGTGCGCTC CGACGGGATG CGGACGTTGA AACCCCACTC CCTCGACTGC TGCACGTTGA
301    AGCTGAAGTC GGGCTTCGGC CGCAGGAAGA AGACGCTCGA CGTGCGGCTG CTGCCGACGC
361    GCTGGTCGCG CCCGGGAGAC TTGACCCGGT GCTTGGGCGC CTTGACCTTG CCGCCCGTGG
421    CCAGGTGTCG GACCGCGCCG CAGAAGACGA CCATGGCGCC GGGGAGCGTC GGGAGGTCGA
481    CGAATTCTCC GTCCACCTCG CACTGCAGGC TCACGAAGCC GTTGGCGCAG GCTGTCTGGT
541    GCACGAGCGT GATGGTCGAT AGGTCGTAGT GGGGTCCCAT GCGGAGGGGT TCCTCTTCGG
601    CGACGCGGTC CTCCGGCACT TCGGGGAAGT ACCGTAGGCG GAGGAGGGGA TCGCACTCGA
661    CGAAGTCATC AATGTCTCTC CCGGCGAGCG GGGCGCCAC AGAGTTGAGA ACGGCGCGCG
721    CGACATCCTT GGCTGCGCCG TACATGCGGT CGAAGTAGTC CTGCCAGACG TCCTCGAAGC
781    CCCGGTTCGG GAACAGGTTG CCGCCGATGC CCATGGAGTA GCACGTCGAG TAGTCCGAGT
841    ACTTGCCCGT CTCGGTGACG ACGGCGGTGC TCTCCCACTC GAGGGCAGAG AAGCCGCGGC
901    GGGCGTTACG GTCGGCGAGC GTCACGGCCC TCTTCTCCTC CTCGCTTCG TTCTTGAAAA
961    AGTCAACGCA CGTCTCACGC GCCGAGGTGT GGTCGTCGTC GACCAGGCCG CTCTCGGTCA
1021   AGTAGAAGAT ACCCTTGGTG GTGACGGCCT CGGCGAGCTC GGTGAGGACC TTGCCGCTCT
1081   TGAGGTCGTC GAGACGAAAG ACGGGGACCT TGGAAGTCAT GTTGATGCTG TGGTTTTGAG
1141   CGATGACTTG AGAGGAGTAG CGTGAGGAA AGTTCTGCAA GAGGAATTTA AGGATTCACA
1201   AGATCCCAGT GAGAACGAAA CGTTGTCAAA GCGGTATATA TATATCTCAA ACCCCACCTC
1261   GTAGCTTACG CCGAGGAACT CCTTTTGA CAACTGCTAC TTAGCCGTAA GTGACGCCCT
1321   GCTTCCCCTC AGCCTTGGCC GCACACGTCA ATGTAGCATT GTAAACCCAC GAGTGTCTTG
1381   TGAAGTTTTG TCAACGAATC ATAAGAAGCC ATCGAGTTCT CTTCTCGTTC TTGGTTTCGA
1441   GGAGAATATG TATCGTGCAT GGTCCCTGAT CGTCGAGACC GCCATGGAAT CGTGCAAGCC
1501   TTAATTCTCC GTACAAGCTT CCCCATTCGG ACAAGATTGC GATGATGTGG ATGCGGGCTC
1561   TTTTAATAAG GACCTTCTTA ACCGATGGTC CGAGAGTGCC TAGGACGGGT CCATGTGCAT
1621   ACACGACGGA CCCTCGACCT CCTATTAGGA GCATGAGGGA CGACAAAATG CGAACGACGA
1681   TGCATCAAAA TGCCTGCAA CGTCGAGTTG TGGGCTACTC GCCTTCTGAT TCGCAAGCCC
1741   TCGGCGAGTC CACCTACTAG TAGCTTGGGA ATAAACAGCA AGTTTCGCCG CCAAAGGGC
1801   TGCCCGGCAT CCGATTCGAT GCCATTGTAC ATCAAGTCGG AAATGGTGCT CCGTTTCCCC

```


1861 CTGGGGTGAG AGGGCGAAGG AGTAGTTCGA CCAGTCGCAG CGCACCCAGA GCCGCAGGTT
1921 TTATCGGATG TTGCTTCGAT CCGATCGTAT CCCGCGCGGC CTAGATCTTG CTAATACGAG
1981 TCGGAGAGTT ACTATTCCGG GCTTATGCGG ACGGGCCGCC GCCGTCGATG CCGGCCAAGG
2041 CTTGTCTGTC ATGATAGATG CTGCCGTCGG CCCAAGTGGC CCGTCTAAAG CCGGACCCCT
2101 TTCCCCGAG TCTCTCCCCG ATCCCGCACG GGGCCGTCAC TTTCGCTGCC CTCGCTCCTT
2161 GTCATAACCT ACCTATATTC TCATCCCGGC AAATGCTGCG GGATAGCCTC ACCTACAGCC
2221 ACACGTCGCC CACCATGTCT CCTCAGATCG CCAATCGCTT CGAGGCTTCG CTAGATGCCC
2281 AAGAĈATAGC CAGAAATATCG CTCTTCACAC TGGAATCTGG CGTCATCCTT CGCGATGTAC
2341 CCGTGGCATA CAAATCGTGG GGTCGCATGA ATGTCTCAAG GGATAACTGC GTCATCGTCT
2401 GCCACACCTT GACGAGCAGC GCCCATGTCA CCTCGTGGTG GCCCACACTG TTTGGCCAAG
2461 GCAGGGCTTT CGATACCTCT CGCTACTTCA TCATCTGCCT AAATTATCTC GGGAGCCCCT
2521 TTGGGAGTGC TGGACCATGT TCACCGGACC CCGATGCAGA AGGCCAGCGC CCGTACGGGG
2581 CCAAGTTTCC TCGCACGACG ATTCGAGATG ATGTTTCGGTA GGTAAGCGCA CCGATCCAGC
2641 TTGTCTCAAT ATCGAGTGGT CAGGACAATC CAGGCTAAGC TTTCCGTGTC CAAAAGTATT
2701 CATCGCCAGG TGCTCGACAG GTTAGGCGTC AGGCAAATTG CTGCCGTAGT CGGCGCATCC
2761 ATGGGTGGAA TGCACACTCT GGAATGGGCC TTCTTTGGTC CCGAGTACGT GCGAAAGATT
2821 GTGCCCATCG CGACATCATG CCGTCAGAGC GGCTGGTGCG CAGCTTGGTT CGAGACACAG
2881 AGGCAGTGCA TCTATGATGA CCCCAGTAC CTGGACGGGG AGTACGACGT AGACGACCAG
2941 CCTGTCCGGG GGCTCGAAAC AGCGCGCAAG ATTGCGAATC TCACGTACAA GAGCAAACCT
3001 GCGATGGACG AGCGCTTCCA TATGGCTCCA GGAGTCCAAG CCGGTGAGTT TATAGATGCC
3061 TTGCCGTGCG TCGATGCTCA GAGCTAATCA GACCGAACCC GCTGCTAGGC CGGAATATCA
3121 GCAGCCAGGA TGCGAAGAAG GAAATCAACG GCACAGACAG CGGCAACAGC CACCGTGCTG
3181 GCCAGCCCAT TGAAGCCGTA TCTTCCTATC TCCGGTACCA GGCCAGAAAG TTTGCCGCGA
3241 GCTTCGACGC CAACTGCTAC ATCGCCATGA CACTCAAGTT CGACACCCAC GACATCAGCA
3301 GAGGCCGGGC AGGATCAATC CCGGAGGCTC TGGCAATGAT TACACAACCA GCGTTGATCA
3361 TTTGCGCCAG GTCAGACGGT CTGTACTCGT TTGACGAGCA CGTTGAGATG GGGCGCAGTA
3421 TCCCAAACAG TCGTCTTTGC GTGGTGGACA CGAATGAGGG TCATGACTTC TTTGTAATGG
3481 AAGCGGACAA GGTTAATGAT GCCGTCAGAG GATTCCCTCGA TCAGTCATTA ATGTGAGGCT
3541 ATGGAGGTGT CAGCCTGCCG GTGCGCGTAC TTGCCAGGGT GATCGATGTA CTCTCAGATA
3601 GTCTCCATGT GAGTATGGAT TTCGCTGTTT CCGCTCGGAT ATAGGCACTC TCAGGCCATC
3661 TCGCAGTAGG TATCAGAACA GCAGCTGAGG CTTTCTCGGA AAGTAGGTTG TGTCAATAGA
3721 TTCATAAAGC GTCAAATAAA GCCCAAAGTC GCAGTAGACT CATCGCATCG CAAGTCTCAG
3781 AGGGTCGACT CGGCAGATTC GAGGCATTGT AGCACATTGT CGAGGCATTG AGGCGGAGAC
3841 TTGACCCATC CAACTCGGCC AGAGGCAGCA GGCAAAGCAT CTCAGCGTAG GCTCCATGCA
3901 AAACATGCGT GGCTCAACTC AGCAAGCTCA TTGCCAACGA GGTCAAAGAA AATAGAAGGT
3961 AGCGGAGGCA GCGGGGTATC GTAGTAACAC CGTCCACATA ACACGGGCTC AGCGGAGCAA
4021 CGTAGTACCT ACTCGTATAG AGGCACCGCG TCAGGAGAGG TATCAGAACC CTCATGATTC
4081 TATCGCCATG CTGCTGCGAA CACTAACAAA TGATAAACAA GGGCCCATGC TGTGTGATGA
4141 TGATTCAAGC AGGTGTGCTG GGTCCAGGTT TGGTGCCCGA GCGCCGCACA GCTGAAGATG
4201 ACGCGTCTCG CTGTCGCGCC TTCCACGACC CAGAAGTTGA TGTGCAGAAT GGGCAGTGAG
4261 TGAACCTGGG CGGGAGTGAT GGAAGGTGCC TACCCTGTAC AACCAACTAC GTCGGTACTC

4321 GTAGGAGCAA TAGCGATGAA GCGTCGGGAG AGAAGTGTGA ATTACTCTGG TACCTGGTAC
4381 TTGATGCAAC ATAGCACATT TCACCCATCA AAGCTAGGTC CCGCGGCCTG GGAGTGGAAT
4441 GGTGAAAGAC ACCGAGGCAA ATGCGGCATG AATGAGGAAG CACGGACGAG TCGTG GTTTC
4501 ACAAGAGACA CTCTGACCGA CCACAAGATT CGGCAGTACA GTCACAGCAT CACCATCGGC
4561 AGTCAGACAT GATTGAGAGC CAGGTCTTCG GCAGAGGGAA TTAGATACAC CTCGGCACCG
4621 GCG